

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 06:11:10 ; Search time 909.231 Seconds  
(without alignments)  
1065.851 Million cell updates/sec

Title: US-10-009-980B-5  
Perfect score: 20  
Sequence: 1 tccaaacaagatggatcg 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Lifting first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hgt.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_phi.\*  
8: gb\_pl.\*  
9: gb\_px.\*  
10: gb\_ro.\*  
11: gb\_sta.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	100.0	405	8	CTE245577
C 2	20	100.0	763	8	AF274687
C 3	19	95.0	220430	2	AC095315
C 4	17.4	87.0	816	8	CNS01CNU
C 5	17.4	87.0	45544	9	AL672201
C 6	17.4	87.0	122394	9	AC011364
C 7	17.4	87.0	169229	2	AC145013
C 8	17.4	87.0	179075	2	AC102895
C 9	17.4	87.0	182906	2	AC131221
C 10	17.4	87.0	186125	10	AC132918
C 11	17.4	87.0	189211	10	AC102745
C 12	17.4	87.0	222978	2	AC124831
C 13	17.4	87.0	226839	2	AC150579
C 14	17.4	87.0	238007	2	AC107569
C 15	17.4	87.0	252660	2	AC138176
C 16	17	85.0	278	6	CQ465431
C 17	17	85.0	138070	2	AC109783
C 18	17	85.0	177056	2	BX950204
C 19	17	85.0	192814	10	AC127417

C	20	17	85.0	214318	10	AC129595
C	21	17	85.0	244978	2	AC128734
C	22	16.8	84.0	1041	8	AF082024
C	23	16.8	84.0	1452	6	AX146368
C	24	16.8	84.0	16775	1	AB000735
C	25	16.8	84.0	64258	8	BX842595
C	26	16.8	84.0	89309	2	AC132864
C	27	16.8	84.0	89309	2	AC132864
C	28	16.8	84.0	110000	2	AC110355_1
C	29	16.8	84.0	110000	2	AC110355_2
C	30	16.8	84.0	110000	8	AB016817_12
C	31	16.8	84.0	110000	8	AB016817_03
C	32	16.8	84.0	125992	5	BX649211
C	33	16.8	84.0	134687	2	AC147472
C	34	16.8	84.0	136886	9	AL138889
C	35	16.8	84.0	161692	9	AC148693
C	36	16.8	84.0	179961	9	AC080166
C	37	16.8	84.0	182853	9	AC074269
C	38	16.8	84.0	186951	2	AC102485
C	39	16.8	84.0	210320	9	AC093873
C	40	16.8	84.0	218329	3	CNS07EGB
C	41	16.8	84.0	224026	10	AC095402
C	42	16.8	84.0	227441	2	CR589875
C	43	16.8	84.0	227441	2	CR589875
C	44	16.8	84.0	239050	1	AL596169
C	45	16.8	84.0	260526	2	AC095888

#### ALIGNMENTS

RESULT 1

LOCUS CTE245577/c

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CTE245577 405 bp DNA linear PLN 06-AUG-2002  
Cyamopsis tetragonoloba partial 5.8S rRNA gene and ITS 2, isolate

seed 10.

AJ245577.1 GI:22138754

26S ribosomal RNA; 26S rRNA gene; 5.8S ribosomal RNA; 5.8S rRNA

gene; internal transcribed spacer 2; ITS2.

Cyamopsis tetragonoloba (guar)

Cyamopsis tetragonoloba

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;

Indigoferaceae; Cyamopsis.

1 Domenech-Sanchez, A., Hernandez, M.L., Rossello, J.A. and Benedi, V.J.

Method for detecting additions of guar gum to locust bean gum

Unpublished

2 (bases 1 to 405)

Domenech-Sanchez, A.

Direct Submission

Submitted (05-AUG-1999) Domenech-Sanchez A., Biologia, Area de

Microbiologia, Universitat de les Illes Balears and IMEDBA

(CSIC-UIB), Carretera de Valldemossa, km 7,5; Palma de Mallorca,

07071, SPAIN

location/Qualifiers

1. .405

/organism="Cyamopsis tetragonoloba"

/mol\_type="genomic DNA"

/isolate="seed 10"

/db\_xref="taxon:3832"

/tissue\_type="seed"

/country="India"

1. .129

/gene="5.8S rRNA"

<1. .129

/gene="5.8S rRNA"

/product="5.8S ribosomal RNA"

130. .346

/note="internal transcribed spacer 2, ITS2"

347. .405

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 05:41:19 ; Search time 220.513 Seconds  
(without alignment)  
536.907 Million cell updates/sec

Title: US-10-009-980b-5

Perfect score: 20

Sequence: 1 tccaacaagaatggagctcg 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001as:\*
- 5: Geneseq2001bs:\*
- 6: Geneseq2002as:\*
- 7: Geneseq2002bs:\*
- 8: Geneseq2003as:\*
- 9: Geneseq2003bs:\*
- 10: Geneseq2003cs:\*
- 11: Geneseq2003ds:\*
- 12: Geneseq2004as:\*
- 13: Geneseq2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	AAH44594	Aah44594 Guar and
2	20	100.0	405	AAH44598	Aah44598 Guar nucl
3	17	85.0	278	ABL85231	Ab185231 Human ova
4	16.8	84.0	1452	AAO50172	Aao50172 Taurus cus
5	16.8	84.0	110000	ABQ69245_16	Continuation (17 o
6	16.8	84.0	110000	ABQ87195_0	Abq87195 Listeria
7	16.4	82.0	187	ADJ67333	Adj67333 Human ova
8	16.4	82.0	599	ADM72157	Adm72157 Human NTR
9	16.4	82.0	1431	AAA38955	Aaa38955 Human G p
10	16.4	82.0	15456	AAC85008	Aac85008 Nucleotid
11	16.4	82.0	15456	AAC85009	Aac85009 Nucleotid
12	16.4	82.0	15456	AAD27765	Aad27765 Bovine pa
13	16.4	82.0	15456	AAD27766	Aad27766 Bovine pa
14	16.4	82.0	52479	ADA02795	Ada02795 Mouse Tnf
15	16.4	82.0	52479	ADB72533	Adb72533 Mouse Tnf
16	16.4	82.0	52479	ADC85275	Adc85275 Mouse Tnf
17	16.4	82.0	52479	ADM74390	Adm74390 Murine ca
18	16	80.0	540	AAA81688	Aaa81688 N. mening
19	16	80.0	49088	ACN43856	Acn43856 Mouse gen
20	15.8	79.0	269	ABL87432	Ab187432 Human ova

21	15.8	79.0	280	12	ADI42908	Adi42908 Plant tra
22	15.8	79.0	280	12	ADO03087	Ado03087 Corn orth
C	23	15.8	392	5	AAS64572	Aas64572 DNA encod
C	24	15.8	400	2	AAV78107	Aav78107 Staphyloc
C	25	15.8	401	4	AAK95811	Aak95811 Human neu
C	26	15.8	401	4	AAK97304	Aak97304 Human neu
C	27	15.8	401	6	ABT00581	Abt00581 Human neu
C	28	15.8	401	6	ABT02074	Abt02074 Human neu
C	29	15.8	431	3	AAC10517	Aac10517 Human sec
C	30	15.8	457	9	ACH47452	Ach47452 Human inf
C	31	15.8	503	6	ABL78849	Ab178849 Human ova
C	32	15.8	553	4	AAI17661	Adi17661 Probe #75
C	33	15.8	553	4	ABA62599	Ada62599 Human foe
C	34	15.8	553	4	AAI42592	Adi42592 Probe #11
C	35	15.8	553	4	ABA29922	Abi29922 Probe #83
C	36	15.8	553	4	AAK36807	Aak36807 Human bon
C	37	15.8	553	4	AAK10959	Aak10959 Human bra
C	38	15.8	553	4	AB36475	Ab36475 Human liv
C	39	15.8	553	6	ABS10811	Abi10811 Human gen
C	40	15.8	561	4	AAI20007	Aai20007 Probe #99
41	15.8	79.0	561	4	ABA65035	Abi65035 Human foe
42	15.8	79.0	561	4	AAI45205	Aai45205 Probe #13
43	15.8	79.0	561	4	ABA47151	Abi47151 Human bre
44	15.8	79.0	561	4	ABA32143	Abi32143 Probe #10
45	15.8	79.0	561	4	AAK39197	Aak39197 Human bon

## ALIGNMENTS

RESULT 1  
AAH44594

ID AAH44594 standard; DNA; 20 BP.

XX AAH44594;

DT 01-NOV-2001 (first entry)

XX Guar and locust bean seed differentiation PCR primer PG21.

XX Guar gum; locust bean gum; detection; plant; initiator; amplification;  
XX PCR; Cyamopsis tetragonoloba; Ceratonia siliqua; thickener;  
XX Gelling agent; food stabiliser; differentiation; PCR primer; ss.

OS Synthetic.

XX WO200166794-A1.

PN 13-SEP-2001.

PD 02-MAR-2001; 2001WO-ES000079.

PF 08-MAR-2000; 2000ES-00000560.

PR (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.

XX (UYIS-) UNIV LAS ISLAS BALEARES.

PA (UYVA-) UNIV VALENCIA.

PA (CARO-) CAROB SA.

XX Benedi Benito VJ, Domenech Sanchez A, Hernandez Viadel Ml;

PI Alberti Serrano S, Rossello Picornell JA;

XX WPI; 2001-565598/63.

XX Differentiating between guar and locust bean seeds, or derived gums, by  
XX amplifying specific, characteristic regions of ribosomal DNA.

PS Claim 4; Fig 1; 44pp; Spanish.

XX The present invention describes a method for differentiating between  
XX seeds of Cyamopsis tetragonoloba (guar) and Ceratonia siliqua (locust  
XX bean) from differences in rDNA extracted from them. The seeds are  
XX germinated, DNA extracted and amplified by polymerase chain reaction

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 06:13:55 ; Search time 1760 Seconds  
(without alignments)  
432.549 Million cell updates/sec

Title: US-10-009-980b-5

Perfect score: 20

Sequence: 1 tccaaacaagatggagctg 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gsa1:\*

9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	18.4	92.0	941	8	AZ541838 ENTFO93TR
2	18	90.0	374	1	AJ012952 EHEST102
3	18	90.0	464	7	CN199108 TGETZy15
4	18	90.0	527	1	AF348499 AF348499
5	18	90.0	864	8	AZ679167 ENTJP67TR
6	18	90.0	873	8	BH152072 ENTJP94TF
7	18	90.0	888	8	AZ683940 ENTHX10TR
8	18	90.0	962	8	BH151479 ENTQ179TR
9	18	90.0	963	8	BH155570 ENTST28TF
10	17.4	87.0	630	7	CV007480
11	17.4	87.0	653	6	CD306610
12	17.4	87.0	760	9	BX958633 Reverse s
13	17.4	87.0	834	9	CR243660 Reverse s
14	17.4	87.0	1604	1	AL692767
15	17	85.0	278	1	AA847512
16	17	85.0	278	1	AA847520
17	17	85.0	381	7	M75738
18	17	85.0	499	4	BJ563859
19	17	85.0	504	4	BJ554532
20	17	85.0	509	4	BJ562795
21	17	85.0	510	4	BJ565764
22	17	85.0	562	4	BJ559871
23	17	85.0	1421	4	BM556682
24	17	85.0	1906	3	CR703050 Tetraodon

C 25 16.8 84.0 283 2 BE339708  
 26 16.8 84.0 325 5 BP752611  
 27 16.8 84.0 350 9 CG397380  
 28 16.8 84.0 402 8 AZ195266  
 29 16.8 84.0 443 2 AQ184987  
 30 16.8 84.0 490 2 BE444732  
 31 16.8 84.0 509 2 BE446001  
 32 16.8 84.0 513 7 CN856183  
 33 16.8 84.0 528 2 BE443711  
 34 16.8 84.0 529 2 BE444665  
 35 16.8 84.0 529 8 AQ378083  
 36 16.8 84.0 529 8 AQ511337  
 37 16.8 84.0 555 6 CD452707  
 38 16.8 84.0 566 7 CF367946  
 39 16.8 84.0 677 6 CA115075  
 40 16.8 84.0 677 9 CE675048  
 41 16.8 84.0 678 5 BQ514206  
 42 16.8 84.0 685 8 BH095996  
 43 16.8 84.0 686 7 CK953792  
 44 16.8 84.0 689 8 BZ001535  
 45 16.8 84.0 689 8 BZ001535

## ALIGNMENTS

RESULT 1

AZ541838/c

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AZ541838 941 bp DNA linear GSS 14-NOV-2000  
 ENTFO93TR Entamoeba histolytica Sheared DNA Entamoeba histolytica  
 genomic, genomic survey sequence.

ACCSSION AZ541838 GI:11149946

VERSION GSS.

KEYWORDS Entamoeba histolytica

SOURCE Entamoeba histolytica

ORGANISM Eukaryota; Entamoebidae; Entamoeba.

REFERENCE 1 (bases 1 to 941)

AUTHORS Loftus B., Van Aken S. and Fraser C.

TITLE Determination of clone end sequences from Entamoeba histolytica

JOURNAL HMI:IMSS sheared DNA library

COMMENT Contact: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: bjloftus@tigr.org

Clones are derived from the Entamoeba histolytica HMI:IMSS sheared

DNA library

Seq primer: M13-Reverse

Class: shotgun

High quality sequence start: 15

High quality sequence stop: 732.

FEATURES

source

1. .941

/organism="Entamoeba histolytica"

/mol\_type="genomic DNA"

/strain="HMI:IMSS"

/db\_xref="taxon:5759"

/clone\_lib="Entamoeba histolytica Sheared DNA"

/note="Vector: PHOSI; Site 1: Bst I; Constructed at The

Institute for Genomic Research (TIGR), Rockville, MD.

Genomic DNA isolated from broth cultures of E. histolytica

using a method described by Clark and Diamond (Clark,

C.G. and Diamond, L.S. (1993) Entamoeba histolytica: a

method for isolate identification. Exp. Parasitol.

77:450.). The DNA was mechanically sheared to give a

tight size distribution (~2 kb). The v + i method used for

the library construction is described in detail in Smith,

H.O. and Venter, J.C. (Making small insert libraries for

whole genome shotgun sequencing projects. In Genome

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 07:10:45 ; Search time 74.8718 Seconds  
(without alignments)  
437.088 Million cell updates/sec

Title: US-10-009-980B-5

Perfect score: 20

Sequence: 1 tccaaacagatggagtcg 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*

2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*

3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*

4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*

5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*

6: /cgn2\_6/ptodata/1/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	16.8	84.0	601	4	US-09-949-016-36311, A
C 2	16.8	84.0	601	4	US-09-949-016-36312, A
C 3	16.8	84.0	601	4	US-09-949-016-36313, A
C 4	16.8	84.0	601	4	US-09-949-016-86588, A
C 5	16.8	84.0	601	4	US-09-949-016-86589, A
C 6	16.8	84.0	601	4	US-09-949-016-86590, A
C 7	16.8	84.0	601	4	US-09-949-016-86591, A
C 8	16.8	84.0	601	4	US-09-949-016-163281, A
C 9	16.8	84.0	601	4	US-09-949-016-163282, A
C 10	16.8	84.0	601	4	US-09-949-016-163283, A
C 11	16.8	84.0	1452	4	US-10-142-231-49, A
C 12	16.8	84.0	119930	4	US-09-949-016-12677, A
C 13	16.8	84.0	119931	4	US-09-949-016-16319, A
C 14	16.4	82.0	601	4	US-09-949-016-61297, A
C 15	16.4	82.0	601	4	US-09-949-016-61298, A
C 16	16.4	82.0	601	4	US-09-949-016-61299, A
C 17	16.4	82.0	601	4	US-09-949-016-67735, A
C 18	16.4	82.0	601	4	US-09-949-016-67736, A
C 19	16.4	82.0	601	4	US-09-949-016-67737, A
C 20	16.4	82.0	47199	4	US-09-949-016-12144, A
C 21	16.4	82.0	47200	4	US-09-949-016-13526, A
C 22	16.4	82.0	163181	4	US-09-949-016-13730, A
C 23	16.4	82.0	268449	4	US-09-949-016-17244, A
C 24	16.4	82.0	17593	4	US-09-949-016-12145, A
C 25	16.4	82.0	17593	4	US-09-949-016-15734, A
C 26	15.8	79.0	400	4	US-08-956-1718-3796, A
C 27	15.8	79.0	400	4	US-08-781-986A-3796, A
C 28	15.8	79.0	410	4	US-09-270-767-6237, A

## ALIGNMENTS

### RESULT 1

US-09-949-016-36311/c

; Sequence 36311, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CLO01307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 36311

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-36311

Query Match 84.0%; Score 16.8; DB 4; Length 601;

Best Local Similarity 90.0%; Pred. No. 1e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCCAAACAGATGGAGTCG 20

Db 41 TCCTAAACAGATGGAGTCG 22

### RESULT 2

US-09-949-016-36312/c

; Sequence 36312, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CLO01307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

Sequence 21519, A  
Sequence 14592, A  
Sequence 945, App  
Sequence 24, Appl  
Sequence 193, App  
Sequence 191, App  
Sequence 191, App  
Sequence 15873, A  
Sequence 14432, A  
Sequence 14433, A  
Sequence 12609, A  
Sequence 17567, A  
Sequence 16246, A  
Sequence 22355, A  
Sequence 25584, A  
Sequence 193956, A  
Sequence 6018, Ap  
Sequence 12126, A

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 08:35:46 ; Search time 339.487 Seconds  
(without alignments)  
385.812 Million cell updates/sec

Title: US-10-009-980B-5  
Perfect score: 20  
Sequence: 1 tccaaacaagatggatcg 20

Scoring table:  
IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7338684 seqs, 3274456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	20	100.0	20	18	US-10-009-980B-5
2	20	100.0	405	18	US-10-009-980B-9
3	17	85.0	278	9	US-09-867-701-8209
4	16.8	84.0	745	13	US-10-027-632-14575
5	16.8	84.0	745	13	US-10-027-632-14576
6	16.8	84.0	745	13	US-10-027-632-14577
7	16.8	84.0	745	13	US-10-027-632-14578

8	16.8	84.0	745	17	US-10-027-632-14575	Sequence 14575, A
9	16.8	84.0	745	17	US-10-027-632-14576	Sequence 14576, A
10	16.8	84.0	745	17	US-10-027-632-14577	Sequence 14577, A
11	16.8	84.0	745	17	US-10-027-632-14578	Sequence 14578, A
12	16.8	84.0	1452	14	US-10-142-231-49	Sequence 49, Appl
13	16.8	84.0	1452	16	US-10-356-153-49	Sequence 49, Appl
14	16.8	84.0	1452	20	US-10-884-115-49	Sequence 49, Appl
15	16.8	84.0	495269	17	US-10-398-221-8	Sequence 8, Appl
16	16.8	84.0	3011208	17	US-10-398-221-2058	Sequence 2058, Ap
17	16.4	82.0	25	21	US-10-719-900-612681	Sequence 612681,
18	16.4	82.0	15456	10	US-09-900-112-35	Sequence 35, Appl
19	16.4	82.0	15456	10	US-09-900-112-36	Sequence 35, Appl
20	16.4	82.0	52479	11	US-09-997-722-61	Sequence 61, Appl
21	16.4	82.0	1790242	20	US-10-719-993-6940	Sequence 6940, Ap
22	16	80.0	422	19	US-10-437-963-13717	Sequence 13717, A
23	16	80.0	466	19	US-10-674-124A-16055	Sequence 16055, A
24	16	80.0	608	20	US-10-425-115-34769	Sequence 34769, A
25	16	80.0	670	20	US-10-425-115-143096	Sequence 143096,
26	16	80.0	49088	13	US-10-087-192-13	Sequence 13, Appl
27	16	80.0	56773	19	US-10-331-053-47	Sequence 47, Appl
28	16	80.0	154817	17	US-10-085-117-334	Sequence 334, App
29	15.8	79.0	269	9	US-09-867-701-10410	Sequence 10410, A
30	15.8	79.0	280	17	US-10-374-780A-1371	Sequence 1371, Ap
31	15.8	79.0	280	18	US-10-412-699B-1500	Sequence 1500, Ap
32	15.8	79.0	379	18	US-10-424-599-106566	Sequence 106566,
33	15.8	79.0	400	8	US-08-781-986A-3796	Sequence 3796, Ap
34	15.8	79.0	400	18	US-10-329-624-3796	Sequence 3796, Ap
35	15.8	79.0	401	9	US-09-795-668-610	Sequence 610, App
36	15.8	79.0	401	9	US-09-795-668-610	Sequence 610, App
37	15.8	79.0	401	9	US-09-946-807-610	Sequence 610, App
38	15.8	79.0	457	10	US-09-918-995-34664	Sequence 34664, A
39	15.8	79.0	503	9	US-09-867-701-1827	Sequence 1827, Ap
40	15.8	79.0	553	9	US-09-864-761-8388	Sequence 8388, Ap
41	15.8	79.0	556	20	US-10-425-115-177887	Sequence 177887,
42	15.8	79.0	561	9	US-09-864-761-17463	Sequence 17463, A
43	15.8	79.0	600	22	US-10-972-079-11643	Sequence 11643, A
44	15.8	79.0	684	21	US-10-487-901-2193	Sequence 2193, Ap
45	15.8	79.0	969	9	US-09-864-761-9173	Sequence 9173, Ap

ALIGNMENTS

RESULT 1  
US-10-009-980B-5  
; Sequence 5, Application US/10009980B  
; Publication No. US20040072155A1  
; GENERAL INFORMATION:  
; APPLICANT: CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS  
; TITLE OF INVENTION: MOLECULAR METHODS FOR DETECTING GUAR GUM ADDITIONS  
; TITLE OF INVENTION: TO LOCUST BEAN GUM  
; FILE REFERENCE: PATENT APPLICATION PCT/ES01/00079  
; CURRENT APPLICATION NUMBER: US/10/009,980B  
; PRIOR FILING DATE: 2002-12-17  
; PRIOR APPLICATION NUMBER: ES200000560  
; PRIOR FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of the Artificial Sequence:oligo ITS2  
US-10-009-980B-5

Query Match	100.0%;	Score 20;	DB 18;	Length 20;
Best Local Similarity	100.0%;	Pred. No. 5.6;		
Matches	20;	Conservative	0;	Mismatches 0;
		Indels	0;	Gaps 0;
Qy	1	TCCAAAACAAGATGGATCG	20	
Db	1	TCCAAAACAAGATGGATCG	20	

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 06:11:10 ; Search time 863.769 Seconds  
(without alignments)  
1065.851 Million cell updates/sec

Title: US-10-009-980B-6

Perfect score: 19  
Sequence: 1 tgcctggcgctgcgcgcgc 19

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_ets.\*  
12: gb\_ey.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	100.0	405	8	CTE245577
2	19	100.0	763	8	AF274687 Cyamopsis
3	17.4	91.6	173	8	AF482510 Utricular
4	17.4	91.6	599	8	AF187091 Pterodon
5	17.4	91.6	602	8	MRU492617 Mitragyna
6	17.4	91.6	613	8	AF086828 Arbutus m
7	17.4	91.6	618	8	AY348776 Impatiens
8	17.4	91.6	620	8	AF467041 Cyclobi
9	17.4	91.6	625	8	AY429098 Oparanthu
10	17.4	91.6	637	8	AF091957 Arbutus x
11	17.4	91.6	639	8	AF091958 Arbutus a
12	17.4	91.6	640	8	AF091956 Arbutus p
13	17.4	91.6	641	8	AF091955 Arbutus o
14	17.4	91.6	642	8	AF061889 Dubautia
15	17.4	91.6	642	8	AF061890 Dubautia
16	17.4	91.6	642	8	AF061891 Dubautia
17	17.4	91.6	642	8	AF061893 Dubautia
18	17.4	91.6	642	8	AF061894 Dubautia
19	17.4	91.6	642	8	AF061896 Dubautia

20	17.4	91.6	642	8	AF061903 Dubautia
21	17.4	91.6	642	8	AF061904 Dubautia
22	17.4	91.6	642	8	AF061905 Dubautia
23	17.4	91.6	642	8	AF061912 Dubautia
24	17.4	91.6	643	8	AF465865 Ichthyoch
25	17.4	91.6	643	8	AF465865 Ichthyoch
26	17.4	91.6	644	8	AF331961 Bidens co
27	17.4	91.6	644	8	AF331962 Bidens co
28	17.4	91.6	719	8	AY634977 Polyplep
29	17.4	91.6	300363	1	AE016781 Pseudomon
30	17.4	91.6	302998	1	AE016921 Chromobac
31	17	89.5	608	8	AY209333 Orobanch
32	17	89.5	642	8	AF061899 Dubautia
33	16.4	86.3	318	8	AF412888 Barnadesi
34	16.4	86.3	436	6	IO1733 Sequence 3
35	16.4	86.3	528	8	AF315496 Camellia
36	16.4	86.3	585	1	RMN1FKDH2
37	16.4	86.3	607	8	AF531087 Brunnera
38	16.4	86.3	622	8	AF456266 Pyrenaria
39	16.4	86.3	624	8	AF229259 Blepharis
40	16.4	86.3	626	8	AY034853 Hebe macr
41	16.4	86.3	630	8	AF456277 Tutcheria
42	16.4	86.3	630	8	AF456280 Tutcheria
43	16.4	86.3	631	8	AF456281 Tutcheria
44	16.4	86.3	634	8	AF456273 Tutcheria
45	16.4	86.3	642	8	AF456267 Pyrenaria

## ALIGNMENTS

RESULT 1	CTE245577	405 bp	DNA	linear	PLN 06-AUG-2002
LOCUS	Cyamopsis tetragonoloba partial 5.8S rRNA gene and ITS 2, isolate				
DEFINITION	seed 10.				
ACCESSION	AJ245577				
VERSION	AJ245577.1	GI:22138754			
KEYWORDS	26S ribosomal RNA; 26S rRNA gene; 5.8S ribosomal RNA; 5.8S rRNA gene; internal transcribed spacer 2; ITS2.				
SOURCE	Cyamopsis tetragonoloba (guar)				
ORGANISM	Cyamopsis tetragonoloba				
REFERENCE	1				
AUTHORS	Domenech-Sanchez, A., Hernandez, M.L., Rossello, J.A. and Benedi, V.J.				
TITLE	Method for detecting additions of guar gum to locust bean gum				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 405)				
AUTHORS	Domenech-Sanchez, A.				
TITLE	Direct Submission				
JOURNAL	Submitted (05-AUG-1999) Domenech-Sanchez A., Biologia, Area de Microbiologia, Universitat de les Illes Balears and IMEDEA (CSIC-UIB), Carretera de Valldemossa, km 7,5; Palma de Mallorca, 07071, SPAIN				
FEATURES	Location/Qualifiers				
source	1..405				
	/organism="Cyamopsis tetragonoloba"				
	/mol_type="genomic DNA"				
	/isolate="seed.10"				
	/db_xref="taxon:3832"				
	/tissue_type="seed"				
gene	1..129				
	/country="India"				
rRNA	/gene="5.8S rRNA"				
	<1..129				
	/gene="5.8S rRNA"				
	/product="5.8S ribosomal RNA"				
misc_feature	130..346				
	/note="internal transcribed spacer 2, ITS2"				
gene	347..405				

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 05:41:19 ; Search time 209.487 Seconds  
(without alignments)  
536.907 Million cell updates/sec

Title: US-10-009-980B-6  
Perfect score: 19  
Sequence: 1 tgcctggcgtcgccgtc 19

Scoring table:  
IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04.\*  
1: Geneseq1980s.\*  
2: Geneseq1980s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001as.\*  
5: Geneseq2001bs.\*  
6: Geneseq2002as.\*  
7: Geneseq2002bs.\*  
8: Geneseq2003as.\*  
9: Geneseq2003bs.\*  
10: Geneseq2003cs.\*  
11: Geneseq2003ds.\*  
12: Geneseq2004as.\*  
13: Geneseq2004bs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	100.0	19	4 AAH4595	AAH4595 Guar and
2	19	100.0	405	4 AAH4598	AAH4598 Guar nucl
C 3	16.4	86.3	654	13 ADT42091	Adt42091 Bacterial
C 4	16.4	86.3	1065	13 ADT48263	Adt48263 Bacterial
5	16.4	86.3	4216	1 AAH70558	Aan70558 Sequence
6	16.4	86.3	110000	2 AAH30458_4	Continuation (5 of
7	16.4	86.3	110000	2 AAH30459_4	Continuation (5 of
C 8	16.4	86.3	110000	4 AAH199682_27	Continuation (28 o
C 9	16.4	86.3	110000	4 AAH199683_27	Continuation (28 o
C 10	15.8	83.2	284	8 ABX45083	Abx45083 Bovine ES
C 11	15.8	83.2	474	11 ABD14195	Abd14195 Pseudomon
C 12	15.8	83.2	708	13 ADS48930	Ads48930 Bacterial
C 13	15.8	83.2	754	6 AAD23833	Aad23833 Tulipa ed
C 14	15.8	83.2	870	2 AAX18428	Aax18428 P. ginsen
C 15	15.8	83.2	870	6 AAD23815	Aad23815 Panax gin
C 16	15.8	83.2	875	2 AAX18432	Aax18432 P. ginsen
C 17	15.8	83.2	875	2 AAX18431	Aax18431 P. trifol
C 18	15.8	83.2	875	2 AAX18427	Aax18427 P. ginsen
C 19	15.8	83.2	875	6 AAD23814	Aad23814 Panax gin
C 20	15.8	83.2	875	6 AAD23818	Aad23818 Panax tri

C 21	15.8	83.2	875	6 AAD24231	Aad24231 Alternati
C 22	15.8	83.2	875	6 AAD23812	Aad23812 Panax qui
C 23	15.8	83.2	1053	6 ADE53368	Ades53368 PEN-1 tel
C 24	15.8	83.2	1053	6 ADE53211	Ades53211 PEN-1 tel
C 25	15.8	83.2	1053	9 ADA66135	Ada66135 DNAP-rela
C 26	15.8	83.2	1620	11 ABD14380	Abd14380 Pseudomon
C 27	15.8	83.2	1632	5 AAH77875	Aah77875 Nucleotid
C 28	15.8	83.2	1632	8 ACA51799	Acas1799 Prokaryot
C 29	15.8	83.2	1800	11 ABD13892	Abd13892 Pseudomon
C 30	15.8	83.2	9006	5 AAH77873	Aah77873 Nucleotid
C 31	15.8	83.2	11658	4 ABL17680	Ab117680 Drosophil
C 32	15.4	81.1	121	6 ABK26702	Abk26702 Waxy star
C 33	15.4	81.1	121	6 ABK26694	Abk26694 Waxy star
C 34	15.4	81.1	121	6 ABK26698	Abk26698 Waxy star
C 35	15.4	81.1	121	6 ABK26697	Abk26697 Waxy star
C 36	15.4	81.1	121	6 ABK26701	Abk26701 Waxy star
C 37	15.4	81.1	121	6 ABK26693	Abk26693 Waxy star
C 38	15.4	81.1	121	12 ADN45385	Adn45385 Mutant ce
C 39	15.4	81.1	121	12 ADN45388	Adn45388 Mutant ce
C 40	15.4	81.1	121	12 ADN45384	Adn45384 Mutant ce
C 41	15.4	81.1	121	12 ADN45389	Adn45389 Mutant ce
C 42	15.4	81.1	121	12 ADN45392	Adn45392 Mutant ce
C 43	15.4	81.1	121	12 ADN45393	Adn45393 Mutant ce
C 44	15.4	81.1	537	11 ABD16439	Abd16439 Pseudomon
C 45	15.4	81.1	723	8 ACA50980	Acas50980 Prokaryot

## ALIGNMENTS

### RESULT 1

AAH44595

ID AAH44595 standard; DNA; 19 BP.

AC AAH44595;

DT 01-NOV-2001 (first entry)

XX XX  
Guar and locust bean seed differentiation PCR primer PG22.

XX XX  
Guar gum; locust bean gum; detection; plant; initiator; amplification;  
PCR; Cyanopsis tetragonoloba; Ceratonia siliqua; thickener;  
XX XX  
Gelling agent; food stabiliser; differentiation; PCR primer; ss.  
XX OS  
Synthetic.  
XX XX  
W0200166794-A1.  
XX PD  
13-SEP-2001.

XX XX  
02-MAR-2001; 2001WO-ES0000079.  
XX PF  
02-MAR-2001; 2000ES-00000560.  
XX PR  
08-MAR-2000; 2000ES-00000560.

XX XX  
(CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
(UYIS-) UNIV LAS ISLAS BALEARES.  
PA PA  
(UYVA-) UNIV VALENCIA.  
PA PA  
(CARO-) CAROB SA.

XX XX  
Benedi Benito VJ, Domenech Sanchez A, Hernandez Viadel ML;  
PI PI  
Alberti Serrano S, Rossello Picornell JA;  
XX XX  
WPI; 2001-565598/63.

XX XX  
Differentiating between guar and locust bean seeds, or derived gums, by  
amplifying specific, characteristic regions of ribosomal DNA.

XX PS  
Claim 4; Fig 1; 44pp; Spanish.

XX CC  
The present invention describes a method for differentiating between  
seeds of Cyamopsis tetragonoloba (guar) and Ceratonia siliqua (locust  
bean) from differences in rDNA extracted from them. The seeds are  
germinated, DNA extracted and amplified by polymerase chain reaction

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 06:13:55 ; Search time 1672 Seconds  
(without alignments)  
432.549 Million cell updates/sec

Title: US-10-009-980b-6

Perfect score: 19  
Sequence: 1 tgcctggcgctgcgcgtc 19

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.\*

1: gb\_est1.\*

2: gb\_est2.\*

3: gb\_est3.\*

4: gb\_est4.\*

5: gb\_est5.\*

6: gb\_est6.\*

7: gb\_est7.\*

8: gb\_gss1.\*

9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17.4	91.6	699	8	CC117786
2	17.4	91.6	712	8	CC129374
3	17.4	91.6	786	9	CC852692
4	17.4	91.6	805	8	CC075421
5	17.4	91.6	840	8	CC080902
6	17.4	91.6	882	8	CC071397
7	17.4	91.6	905	9	CG381256
8	17.4	91.6	917	8	CC101845
9	17.4	91.6	939	9	CG381268
10	17.4	91.6	940	9	CG084378
11	17.4	91.6	998	9	CG084375
12	17.4	89.5	996	2	BE739997
13	16.4	86.3	332	7	CC101845
14	16.4	86.3	357	1	AL831066
15	16.4	86.3	366	2	BE419987
16	16.4	86.3	403	1	AI737701
17	16.4	86.3	432	6	CA012158
18	16.4	86.3	454	6	CD893052
19	16.4	86.3	464	6	CA710712
20	16.4	86.3	477	6	CD923546
21	16.4	86.3	504	6	CA400135
22	16.4	86.3	511	1	AU245885
23	16.4	86.3	522	7	CN864876
24	16.4	86.3	529	6	CB858536

C	25	16.4	86.3	532	4	BJ211978
C	26	16.4	86.3	533	1	AI670298
C	27	16.4	86.3	535	2	BE517064
C	28	16.4	86.3	535	7	CN914909
C	29	16.4	86.3	543	7	CN912613
C	30	16.4	86.3	554	6	CD442907
C	31	16.4	86.3	556	1	AI737987
C	32	16.4	86.3	561	1	AI657458
C	33	16.4	86.3	563	5	EX266910
C	34	16.4	86.3	565	5	BQ467498
C	35	16.4	86.3	567	4	BJ304612
C	36	16.4	86.3	567	7	CN894586
C	37	16.4	86.3	574	5	EX266909
C	38	16.4	86.3	580	7	CN948867
C	39	16.4	86.3	590	7	CV084369
C	40	16.4	86.3	591	5	BQ467224
C	41	16.4	86.3	591	7	CN902278
C	42	16.4	86.3	594	6	CA730170
C	43	16.4	86.3	597	4	BI222007
C	44	16.4	86.3	597	5	BP449990
C	45	16.4	86.3	601	7	CN901560

## ALIGNMENTS

RESULT 1  
CC117786  
LOCUS  
DEFINITION  
CC117786 699 bp DNA linear GSS 16-APR-2003  
NDL.31A3.T7 Notre Dame Liverpool Aedes aegypti genomic clone  
NDL.31A3, genomic survey sequence.  
CC117786  
CC117786.1 GI:29986841  
GSS.  
SOURCE  
Aedes aegypti (yellow fever mosquito)  
ORGANISM  
Aedes aegypti  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Aedes;  
Stegomyia.

REFERENCE  
1 (bases 1 to 699)  
Loftus, B., Shetty, J., Knudson, D. and Severson, D.

AUTHORS  
BAC end sequencing of Aedes aegypti

TITLE  
Unpublished (2003)

JOURNAL  
JOURNAL

COMMENT  
Other GSSs: NDL.31A3.SP6

Contact: Brendan Loftus

Department of Eukaryotic Genomics

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-3543

Fax: 301-838-0208

Email: enta@tigr.org

Library was provided by David Severson

Seq primer: T7

Class: BAC ends.

## FEATURES

source

1..699

/organism="Aedes aegypti"

/mol\_type="genomic DNA"

/strain="Liverpool"

/db\_xref="taxon:7159"

/clone="NDL.31A3"

/note="Vector: pECBAC1; Site 1: Hind III; The library was

prepared from whole body tissue of newly hatched L1 larvae

by David Severson at the University of Notre Dame and

Hongbin Zhang"

## ORIGIN

Query Match 91.6%; Score 17.4; DB 8; Length 699;  
Best Local Similarity 94.7%; Pred. No. 8.6e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TGCCTGGCGTGC GCGTC 19



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 07:10:45 ; Search time 71.1282 Seconds  
(without alignments)  
437.088 Million cell updates/sec

Title: US-10-009-980B-6  
Perfect score: 19  
Sequence: 1 tgcctggcgctgcgcgtc 19  
Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.4	86.3	536165	4	US-09-214-808-1
2	16.4	86.3	4403785	3	US-09-103-840A-2
3	16.4	86.3	4411529	3	US-09-103-840A-1
4	16	84.2	582	4	US-09-902-540-3481
5	16	84.2	17188	4	US-09-902-540-1166
6	15.8	83.2	474	4	US-09-252-991A-12799
7	15.8	83.2	754	3	US-09-541-941B-26
8	15.8	83.2	870	2	US-08-778-912A-4
9	15.8	83.2	870	3	US-09-541-941B-4
10	15.8	83.2	875	2	US-08-778-912A-2
11	15.8	83.2	875	2	US-08-778-912A-3
12	15.8	83.2	875	2	US-08-778-912A-7
13	15.8	83.2	875	3	US-09-541-941B-1
14	15.8	83.2	875	3	US-09-541-941B-3
15	15.8	83.2	875	3	US-09-541-941B-7
16	15.8	83.2	1053	4	US-09-940-244-393
17	15.8	83.2	1620	4	US-09-252-991A-12984
18	15.8	83.2	1800	4	US-09-252-991A-12496
19	15.8	83.2	4455	4	US-09-902-540-7119
20	15.8	83.2	4458	4	US-09-902-540-650
21	15.4	81.1	350	3	US-08-651-155B-36
22	15.4	81.1	350	4	US-09-194-036B-36
23	15.4	81.1	537	4	US-09-252-991A-15043
24	15.4	81.1	1014	4	US-09-252-991A-15183
25	15.4	81.1	1017	4	US-09-252-991A-15184
26	15.4	81.1	1239	4	US-09-902-540-7944
27	15.4	81.1	1242	4	US-09-902-540-4840

C	28	15.4	81.1	1437	4	US-09-252-991A-1434	Sequence 1434, Ap
	29	15.4	81.1	1473	4	US-09-252-991A-1584	Sequence 1584, Ap
C	30	15.4	81.1	1818	4	US-09-731-166-3	Sequence 3, Appli
C	31	15.4	81.1	2542	3	US-08-941-445A-6	Sequence 6, Appli
C	32	15.4	81.1	4800	3	US-08-941-445A-4	Sequence 4, Appli
C	33	15.4	81.1	11922	4	US-09-902-540-1063	Sequence 1063, Ap
C	34	15.4	81.1	21758	4	US-09-902-540-1238	Sequence 1238, Ap
	35	15.4	81.1	39949	4	US-09-949-016-14210	Sequence 14210, A
	36	15	78.9	405	4	US-09-252-991A-4225	Sequence 4225, Ap
C	37	15	78.9	843	4	US-09-252-991A-4551	Sequence 4551, Ap
C	38	15	78.9	1128	4	US-09-252-991A-4450	Sequence 4450, Ap
	39	15	78.9	1242	4	US-09-252-991A-4304	Sequence 4304, Ap
	40	15	78.9	1302	4	US-09-252-991A-4157	Sequence 4157, Ap
C	41	15	78.9	3333	4	US-09-902-540-4753	Sequence 4753, Ap
	42	15	78.9	14809	4	US-09-902-540-1032	Sequence 1032, Ap
	43	15	78.9	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	44	15	78.9	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	45	14.8	77.9	390	4	US-09-902-540-7078	Sequence 7078, Ap

ALIGNMENTS

RESULT 1  
US-09-214-808-1  
; Sequence 1, Application US/09214808A  
; Patent No. 6475793  
; GENERAL INFORMATION:  
; APPLICANT: Rosenthal, Andre  
; APPLICANT: Freiberg, Christoph  
; APPLICANT: Perret, Xavier Philippe  
; APPLICANT: Broughton, William John  
; TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic  
; Patent No. 6475793  
; TITLE OF INVENTION: Plasmid  
; FILE REFERENCE: CARP0068  
; CURRENT APPLICATION NUMBER: US/09/214.808A  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: PCT/IB97/00950  
; PRIOR FILING DATE: 1997-07-10  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 536165  
; TYPE: DNA  
; ORGANISM: Rhizobium  
US-09-214-808-1

Query Match 86.3%; Score 16.4; DB 4; Length 536165;  
Best Local Similarity 94.4%; Pred. No. 1.3e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GCCTGGCGTCGCGCGTC 19  
|||  
Db 453718 GCCTGGCGTCGCGCATC 453735

RESULT 2  
US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 08:35:46 ; Search time 322.513 Seconds  
(without alignments)  
385.812 Million cell updates/sec

Title: US-10-009-980B-6

Perfect score: 19  
Sequence: 1 tgcctggcgctgcgcgcgc 19

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7338684 seqs, 3274456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US05\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10J\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	100.0	19	18	US-10-009-980B-6
2	19	100.0	405	18	Sequence 6, Appli
3	18	94.7	1964	19	Sequence 9, Appli
4	17.4	91.6	1646	20	Sequence 12174, A
5	16.4	86.3	370	20	Sequence 53483, A
6	16.4	86.3	603	20	Sequence 39828, A
7	16.4	86.3	624	20	Sequence 57132, A
					Sequence 39345, A

C	8	16.4	86.3	654	17	US-10-369-493-40529	Sequence 40529, A
C	9	16.4	86.3	738	20	US-10-425-115-39179	Sequence 39179, A
C	10	16.4	86.3	925	20	US-10-425-115-174017	Sequence 174017, A
C	11	16.4	86.3	1065	17	US-10-369-493-46701	Sequence 46701, A
C	12	16.4	86.3	1516	20	US-10-425-115-79542	Sequence 79542, A
C	13	16.4	86.3	1527	20	US-10-425-115-18205	Sequence 18205, A
C	14	16.4	86.3	2649	19	US-10-437-963-112	Sequence 112, App
C	15	16.4	86.3	53165	10	US-09-939-964-1	Sequence 1, Appli
C	16	16	84.2	1273	20	US-10-425-115-125894	Sequence 125894, A
C	17	16	84.2	3186778	13	US-10-027-632-174961	Sequence 174961, A
C	18	16	84.2	3186778	17	US-10-027-632-174961	Sequence 174961, A
C	19	15.8	83.2	284	9	US-09-960-352-10248	Sequence 10248, A
C	20	15.8	83.2	468	20	US-10-425-115-164790	Sequence 164790, A
C	21	15.8	83.2	600	22	US-10-972-079-8955	Sequence 8955, Ap
C	22	15.8	83.2	652	20	US-10-425-115-68106	Sequence 68106, A
C	23	15.8	83.2	704	20	US-10-425-115-157962	Sequence 157962, A
C	24	15.8	83.2	708	17	US-10-369-493-27360	Sequence 27360, A
C	25	15.8	83.2	754	9	US-09-897-231-26	Sequence 26, Appli
C	26	15.8	83.2	759	20	US-10-425-115-53318	Sequence 53318, A
C	27	15.8	83.2	807	20	US-10-425-115-137985	Sequence 137985, A
C	28	15.8	83.2	855	20	US-10-425-115-45730	Sequence 45730, A
C	29	15.8	83.2	859	20	US-10-425-115-70138	Sequence 70138, A
C	30	15.8	83.2	870	9	US-09-897-231-4	Sequence 4, Appli
C	31	15.8	83.2	875	9	US-09-897-231-1	Sequence 1, Appli
C	32	15.8	83.2	875	9	US-09-897-231-3	Sequence 3, Appli
C	33	15.8	83.2	875	9	US-09-897-231-7	Sequence 7, Appli
C	34	15.8	83.2	938	20	US-10-425-115-70438	Sequence 70438, A
C	35	15.8	83.2	972	20	US-10-425-115-166184	Sequence 166184, A
C	36	15.8	83.2	996	20	US-10-425-115-56079	Sequence 56079, A
C	37	15.8	83.2	1023	20	US-10-425-115-12243	Sequence 12243, A
C	38	15.8	83.2	1053	10	US-09-940-244-393	Sequence 393, App
C	39	15.8	83.2	1053	12	US-09-732-622A-393	Sequence 393, App
C	40	15.8	83.2	1053	20	US-10-309-584-393	Sequence 393, App
C	41	15.8	83.2	1053	21	US-10-783-557-394	Sequence 394, App
C	42	15.8	83.2	1233	20	US-10-425-115-143856	Sequence 143856, A
C	43	15.8	83.2	1291	20	US-10-425-115-57660	Sequence 57660, A
C	44	15.8	83.2	1553	20	US-10-425-115-118641	Sequence 118641, A
C	45	15.8	83.2	1632	17	US-10-382-122A-39669	Sequence 39669, A

ALIGNMENTS

RESULT 1  
US-10-009-980B-6  
; Sequence 6, Application US/10009980B  
; Publication No. US20040072155A1  
; GENERAL INFORMATION:  
; APPLICANT: CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS  
; TITLE OF INVENTION: MOLECULAR METHODS FOR DETECTING GUAR GUM ADDITIONS  
; FILE REFERENCE: PATENT APPLICATION PCT/ES01/00079  
; CURRENT APPLICATION NUMBER: US/10/009,980B  
; PRIOR FILING DATE: 2002-12-17  
; PRIOR APPLICATION NUMBER: ES200000560  
; PRIOR FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 19  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of the Artificial Sequence:oligo PG22  
US-10-009-980B-6

Query Match	100.0%	Score 19;	DB 18;	Length 19;
Best Local Similarity	100.0%	Pred. No. 17;		
Matches	19;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	TGCCTGGCGCTCGCGCGTC	19	
Db	1	TGCCTGGCGCTCGCGCGTC	19	